Yusuf Çelebi

Applied AI Scientist

İstanbul / Bursa, Türkiye

16celebiyusuf@gmail.com · Phone: +90 545 927 5502

LinkedIn · GitHub · Hugging Face · Website

Summary

Applied AI Scientist working on LLM systems, retrieval, and agentic AI, with experience building production-grade systems by turning research into practical applications. Worked on agentic pipelines, memory systems, and retrieval-based architectures, including applications in legal and financial domains such as compliance and risk analysis. Also involved in academic research with publications including top-tier conference venues.

Experience

Newmind AI — AI Research Engineer — İstanbul, Türkiye

Jul 2025 – Apr 2026

- **LLM fine-tuning** and adaptation pipelines using **PyTorch** and the **Hugging Face** ecosystem (Transformers, Accelerate), with workflows covering supervised fine-tuning, instruction tuning, and parameter-efficient methods (**LoRA**, **QLoRA**, adapter-based approaches) for domain-specific tasks.
- **Agentic RAG** systems and multi-agent workflows using orchestration frameworks (**LangGraph, CrewAI, Microsoft AutoGen**), tool integrations, and structured execution flows for legal-tech applications (litigation, arbitration, regulatory processes).
- Built **multi-agent due diligence workflows** where documents are analyzed by domain-specific expert agents in parallel, producing structured legal risk reports.
- Developed OSINT-based **compliance workflows** that collect external signals and transform them into structured data for **financial and regulatory risk** assessment.
- **LLM evaluation** and robustness framework development, first author of a paper accepted at **MLSys 2026**, focusing on model behavior under persuasion and agreement settings.
- **Graph-based memory and knowledge systems** using **Neo4j** and graph tooling, including **ontology** design, entity linking, and persistent memory architectures for conversational agents.
- Contributed to legal knowledge systems modeling versioned regulations and legislative changes, enabling reasoning over temporal legal context.
- **Retrieval-augmented systems** with dense retrieval (bi-encoders) and late-interaction methods (**ColBERT**), including training pipelines, reranking, and evaluation using BEIR-style frameworks on Turkish datasets (NLI/STS, MS MARCO-TR).
- **Prompt optimization and system-level tuning** using optimization frameworks (**DSPy**, GEPA) for improving retrieval quality, reasoning consistency, and pipeline-level performance.
- **LLM evaluation pipelines** using controlled prompting, log-probability inspection, and structured behavioral testing for robustness and reliability analysis.
- **Data pipelines and automation systems** using scraping and processing stacks, including multi-source extraction, filtering, and XML/structured reporting pipelines
- Built automated pipelines that monitor web and social media sources, perform sentiment analysis, and generate structured risk reports.
- **Representation learning and embedding-space analysis** for transformer models, focusing on geometric structure of hidden states and identifying structurally critical layers for efficient adaptation.
- **End-to-end research integration** across LLMs, retrieval, and agentic systems, owning the full lifecycle from **PoC development** to validated, production-grade deployments.

ISSD Bilişim Elektronik A.Ş. — Computer Vision Engineer Intern — Ankara, Türkiye

Jun 2024 – Oct 2024

- **Computer vision and perception systems** using deep learning-based detection and tracking models (**YOLO**-based pipelines), including real-time video processing and object recognition tasks.

- **End-to-end model development workflows** from data processing to deployment, covering image processing, model integration, and system-level optimization for real-world scenarios.
- **Applied research exposure** to modern vision models (**SAM**), following recent literature and evaluating their applicability in production settings.

Publications

PARROT: Persuasion and Agreement Robustness Rating of Output Truth — A Sycophancy Robustness Benchmark for LLMs
MLSys 2026 (First Author)

TurkColBERT: Benchmarking Dense and Late-Interaction Models for Turkish Information Retrieval
Procedia Computer Science (ACLing 2026)

RDP-LoRA: Geometry-Driven Layer Selection for Parameter-Efficient Adaptation in LLMs
preprint

Technical Skills

Programming: Python, SQL
Machine Learning & Deep Learning: PyTorch, Transformers (Hugging Face), TRL, Unsloth, LLMs, CNNs, model fine-tuning, LoRA / QLoRA / PEFT
LLM & Agentic Systems: RAG, LangChain, LangGraph, multi-agent systems, prompt optimization, evaluation pipelines, MCP
Retrieval & Knowledge Systems: Dense retrieval, ColBERT, reranking, BEIR-style evaluation, Elasticsearch, Neo4j, Cypher
Computer Vision & Data Processing: OpenCV, image processing, YOLO, SAM, data pipelines, web scraping
Tools & Infrastructure: Docker, Git

Education

B.Sc. Computer Engineering Sep 2020 – Jan 2025
Hasan Kalyoncu University

Languages

Turkish (Native)
English (B2)

References

References will be provided upon request.